NPS-53-87-006

# NAVAL POSTGRADUATE SCHOOL

## Monterey, California

SAMPLING INTERVALS

by

Toke Jayachandran

H. J. Larson

Technical Report for Period

October 1986 - September 1987

Approved for public release; distribution unlimited

Prepared for:

mmanding Officer
ly Air Force Base
-ALC/MME1)
as, 79241

NAVAL POSTGRADUATE SCHOOL
MONTEREY CALIFORNIA 93943

R. C. AUSTIN
Rear Admiral, U. S. Navy

K. T. MARSHALL
Acting Academic Dean

The work herein was supported with funds provided by Kelly Air Force Base, (SA-ALC/MME1), Texas, 79241.



This report was prepared by:

# REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

| | |
|---|---|
| 1. REPORT NUMBER: NPS-53-87-006 | 2. GOVT ACCESSION NO. |
| 3. RECIPIENT'S CATALOG NUMBER | |
| 4. TITLE (and Subtitle): SAMPLING INTERVALS | 5. TYPE OF REPORT & PERIOD COVERED: TECHNICAL REPORT 10/86 - 09/87 |
| | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s): T. JAYACHANDRAN, H. J. LARSON | 8. CONTRACT OR GRANT NUMBER(s) |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS: NAVAL POSTGRADUATE SCHOOL MONTEREY, CA 93943 | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: PO MME-87-717A1 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS: Commanding Officer (SA-ALC/MME1) | 12. REPORT DATE |
| | 13. NUMBER OF PAGES: 12 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): KELLY AIR FORCE BASE, TEXAS | 15. SECURITY CLASS. (of this report): UNCLASSIFIED |
| | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

SAMPLING, OIL ANALYSIS, STATISTICAL ANALYSIS, PRINICIPAL COMPONENTS

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

THE STATUS OF SAMPLING INTERVALS BETWEEN OIL ANALYSIS HAS BEEN REVIEWED AND A FRAMEWORK FOR AN ALTERNATIVE APPROACH TO THE DETERMINATION OF SAMPLING INTERVALS IS DISCUSSED.

## SAMPLING INTERVALS

Spectrometric analysis of used oil samples drawn from aircraft engine components is an integral part of the aircraft maintenance program in the United States Air Force. The samples are drawn at prescribed intervals of flight time and analyzed on a spectrometer to monitor the levels of occurrence (in PPM) of certain wear metal contaminants. The observed contaminant levels as well as their rates of growth with time are used to assess the wear condition of an engine and to predict certain types of component failures before they can become critical. The interval between successive samples is called the "sampling interval". These intervals are usually determined at the time of introduction of an aircraft into the fleet, in consultation with the manufacturer. The selected interval is fixed for all aircraft of the same type independent of the age of the aircraft. Typically, single-engine aircraft have shorter sampling intervals and multi-engine aircraft generaly will have longer intervals. Since very little information on the wear metal buildup mechanism would be available for new aircraft the prescribed sampling intervals tend to be conservative; that is, samples are analyzed more frequently than would be necessary. The oil analysis process is quite expensive both in terms of the dollar costs of sample acquisition, transportation, analysis, record maintenance etc., as well as the availability of the aircraft. It would, therefore, be desirable to develop an adaptive scheme that would prescribe a longer sampling interval until the oil analysis indicates the onset of an abnormal wear condition at which time the sampling interval would be shortened, depending on the severity of wear, indicated by the analysis.

The Southwest Research Institute (SRI) conducted a study in 1977 and recommended [5] a new set of sampling intervals for 16 different aircraft/engine types. The statistical methodology applied by SRI is practically the same as the one proposed by ARINC Research Corporation [ 4 ] for the construction of oil analysis decision tables. For each aircraft/engine type, an identification is made of all cases in which an oil analysis resulted in a T-code (ground the aircraft to replace failing parts) recommendation. The results of all the oil analyses, subsequent to the immediately preceding oil change, for each of the identified cases are pooled. A statistical algorithm due to Hudson [1] is employed to fit a semented line (two straight lines jointed at a common point T) to the pooled data. The fitted segmented line is to be the basis for the determination of an appropriate sampling interval for the aircraft type. Two figures (figures A-1 and A-38) taken from the ARINC report are included here for purposes of illustration and discussion. Similar figures showing the fitted segmented lines are also available in the SRI report except that the actual data used is not included in the plots. A basic assumption in adopting this methodology is that wear metal contaminants accumulate linearly with time and at the onset of a malfunction the rate of accumulation shifts to a higher level. The join point T can be thought of as a statistical estimate of the average (over all potential failure mechanisms) time, after an oil change, at which malfunctions are identifiable through oil analysis. SRI's prescription is to choose T/2, T and 3T/2 as the appropriate sampling interval for a single-engine, twin-engine and multi-engine aircraft respectively. On this basis, for the 16 aircraft/engine types included in their study, SRI proposed a new set of sampling intervals that were, in general, smaller than those that were current and in several cases their recommendation would have resulted in a doubling of the sampling frequency. A critique of the SRI approach follows. First, the data consists of the pooled wear metal histories of all aircraft rceiving a T-code, regardless of which failing component (propeller shaft, oil pump, reduction gear box, bearing) caused the issuance of the T-code and also independent of which wear metal(s) exeeded the critical limits. It is reasonable to believe that the amount of change in the rate of accumulation of a wear metal is dependent on the particular component that is failing and also that the set of "significant wear metals" would differ from component to component. If this is the case, pooling oil analysis records over all failures will result in treating several divergent sets of data as a single homogeneous group. It is doubtful that the fitted segmented line would provide an accurate representation of the contaminant growth phenomenon for all future potential failures. Figures A-1 and A-38, we believe, demonstrate this problem. The data on magnesium, plotted in figure A-1, is pooled over two different failure modes (auxillary drive bearing and an oil pump). The plot seems to indicate two distinct groups of data and one of the groups consists of a constant reading of one PPM; perhaps magnesium is not the miscreant wear metal for this group. It is not clear that the fitted segmented line adequately portrays the contaminant growth phenomenon for either group. Figure A-38 shows the data from 15 different failure modes. In this case, the data is so widely dispersed about the fitted lines that it is highly unlikely that the segmented line can be used with any degree of success. A second debatable issue is the idea of a single "wear metal of primary interest" for each aircraft/ engine type, proposed by SRI. Their concept is that for each aircraft/engine type it is possible to select <u>one single wear metal</u> whose wear metal history (ignoring the data on all other wear metals) can be used effectively to monitor the wear status of the engine. In the 35 case histories included in their study they chose either iron, copper or magnesium as the primary wear metal. The implied assumption behind their contention is that all failure modes will <u>always</u> generate excessive amounts of contaminant particles of a pre-specified wear metal. If this were really true the Air Force can save itself a lot of expense by not even monitoring the other (about 10) wear metals. Thirdly, in 21 out of the 35 figures in the SRI report the data consisted of less than 10 case histories. Yet, they recommended the adoption of the sampling intervals derived by them using what appears to be a rather small number of cases. It is true that almost all the newly proposed intervals are shorter than those that were current and hence would not increase the risk of not detecting potential failures in time. But then why change an existing scheme to a potentially more expensive one unless there is evidence to indicate that the current intervals are inadequate. No such evidence is presented in the report. Finally, SRI asserted that their sampling intervals would guarantee a "100-percent probability of obtaining two samples during the abnormal wear period" for a single-engine aircraft. Similar assertions for twin and multi-engine aircraft are also in the report. There is no theoretical or statistical basis for these statements. In fact, no statistical scheme would guarantee 100-percent results.

The Naval Postgraduate School (NPS) in a 1980 technical report [2] discussed the framework for an alternative approach that could lead to more cost effective sampling intervals. This approach is based on the fundamental premise that the wear metal buildup curves for different serial numbers of an aircraft/engine type could be vastly different and hence sampling intervals should be individually tailored. In other words, the contaminant growth characteristics exhibited in the wear metal history for a specific serial number should be the basis for the selection of the most appropriate sampling interval for that unit. One possible approach to the implementation of this scheme is the following. For each serial number, choose an initial sampling interval on an ad hoc basis e.g., twice as long as the currently prescribed interval. After each oil oil analysis, fit a straight line to the wear metal measurements obtained subsequent to the preceding oil change, but excluding the most recent analysis, one for each wear metal. Based on the fitted lines, determine statistical bounds below which the most recent measurements should lie, for a normally functioning engine. If the observed reading for any one of the wear metals exeeds the corrosponding bound shorten the sampling interval to one half of the original length. Otherwise continue sampling at the initial rate. A slightly different approach that will require fitting just one straight line instead one line for each wear metal (thus reducing the necessary computations) is to assume that there exists an "optimal linear combination" of the measurements on the different wear metals that will serve as a "good" discriminant of abnormal wear. There are several ways to estimate such a linear combination. One solution to the estimation of the optimal linear combination is to use a well known statistical technique called the principal components analysis and select the first principal component as the desired linear combination. Once the linear combination is identified all that is needed is to compute its composite value from the results of each of the oil analyses and fit a straight line to these composite scores, excluding the data for the current oil analysis. A statistical upper bound for the composite score for the latest analysis is determined; if this score exeeds the bound, change the sampling interval to one half of the original length. A detailed description of this statistical approach is presented in the appendix.

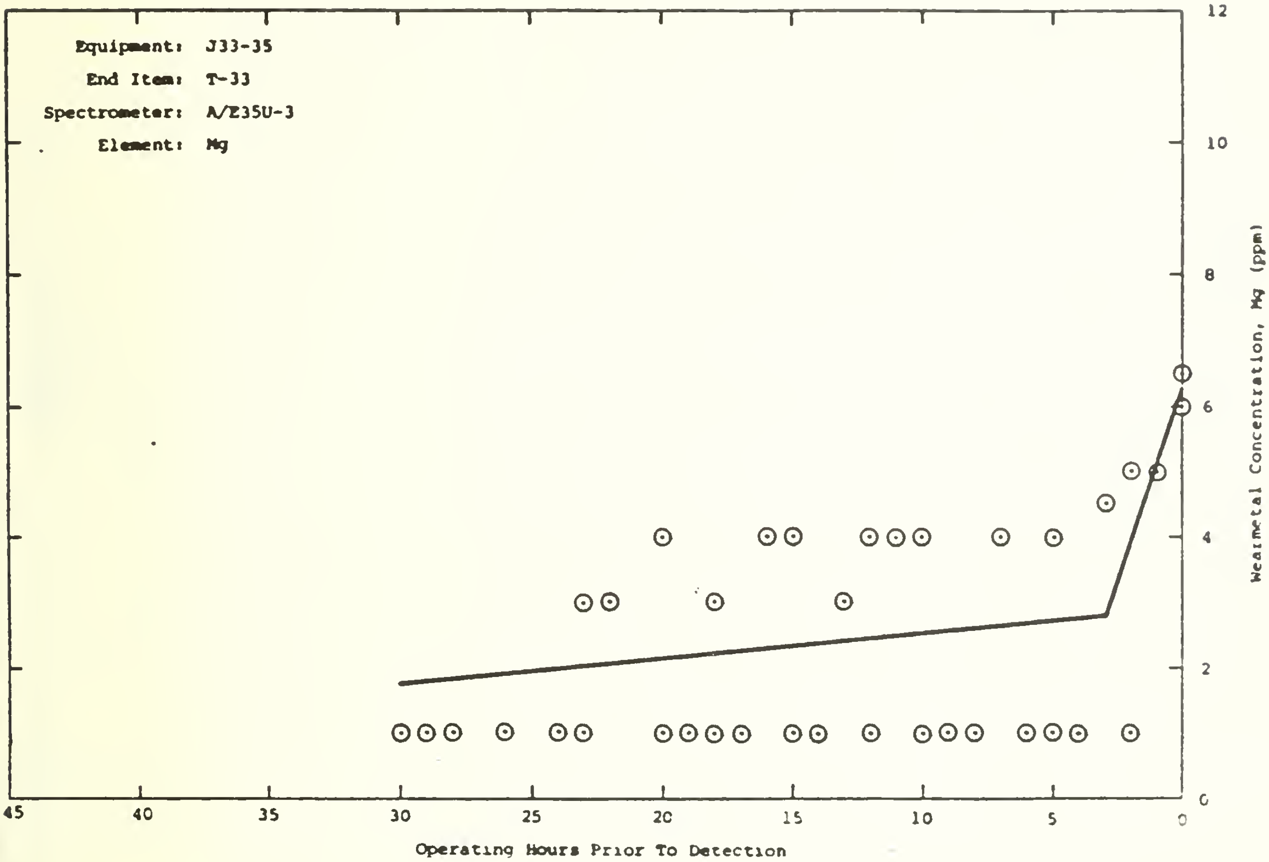
Before this procedure can be considered for adoption, the methodology needs to be tested thoroughly with real data. Some of the questions that need to be answered are:

(1) Is it realistic to assume that a single linear combination an be identified, that will effectively predict all potential failure mechanisms?

(2) How much additional effort on the part of the laboratory personnel would be necessary to make these

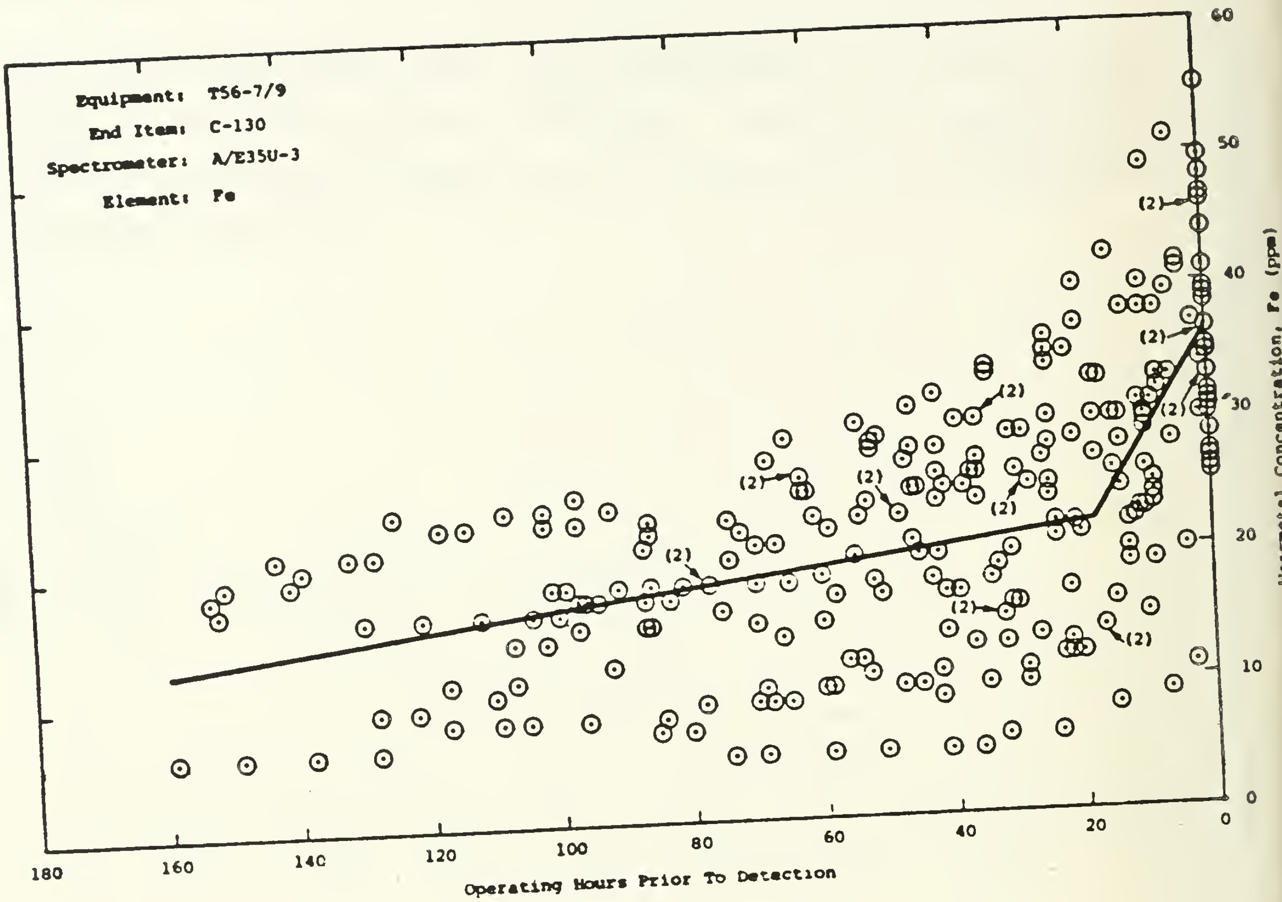
this procedure operational?

(3) Would the adoption of this procedure result in significantly more cost effective sampling intervals without increasing the risk of non-identification of potential failures?




| Equipment S/N | Date of Detection | Reported Malfunction | Equipment S/N | Date of Detection | Reported Malfunction |
|---|---|---|---|---|---|
| 7313 | 720809 | Auxiliary Drives and Accessories, Bearing | | | |
| 7346 | 720620 | Oil Pump, Compressor Section | | | |

*Figure A-1.*




| Equipment S/N | Date of Detection | Reported Malfunction | Equipment S/N | Date of Detection | Reported Malfunction |
|---|---|---|---|---|---|
| 1779 | 720316 | Propeller Shaft, Reduction Gearbox | 3483 | 730602 | Oil Pump, Reduction Gearbox |
| 1973 | 730828 | Oil Pump, Reduction Gearbox | 3889 | 730426 | Oil Pump, Reduction Gearbox |
| 2136 | 730518 | Reduction Gearbox, Bearing | 3889 | 730619 | Oil Pump, Reduction Gearbox |
| 2219 | 730609 | Oil Pump, Power Section | 4183 | 731123 | Reduction Gearbox |
| 2232 | 730516 | Oil Pump, Reduction Gearbox | 4724 | 730921 | Reduction Gearbox, Bearing |
| 2419 | 721201 | Reduction Gearbox, Bearing | 4921 | 720703 | Propeller Shaft, Reduction Gearbox |
| 2620 | 721205 | Oil Pump, Reduction Gearbox | 4921 | 720816 | Reduction Gearbox, Bearing |
| 2626 | 730911 | Oil Pump, Power Section; Scavenge Pump, Power Section, External | 4995 | 731006 | Oil Pump, Reduction Gearbox |
| 2722 | 721110 | Reduction Gearbox, Bearing | 7148 | 730413 | Oil Pump, Reduction Gearbox |
| 2919 | 730910 | Reduction Gearbox | 7148 | 731203 | Scavenge Pump, Rear Turbine |
| 3023 | 720428 | Propeller Shaft, Reduction Gearbox | 7567 | 731004 | Reduction Gearbox |
| 3023 | 730427 | Oil Pump, Reduction Gearbox | | | |
| 3238 | 731121 | Reduction Gearbox | | | |
| 3311 | 720912 | Oil Pump, Reduction Gearbox | | | |
| 3430 | 720706 | Reduction Gearbox | | | |

*Figure A-38.*

References

[1] Hudson, D.J., "Fitting Segmented Curves Whose Join Points Have to Be Estimated", Journal of the American Statistical Association,vol.61, pp 1097-1129, Dec.1966.

[2] Larson, H.J. and Jayachandran, T., " Methodolgy for Determining Sampling Intervals ", Naval Postgraduate School Technical Report NPS-53-81-001, Nov.1980.

[3] Larson, H.J. and Jayachandran, T., " The CEMS IV OAP Algorithm ", Naval Pstgraduate School Technical Report NPS55-83-013, May 1983.

[4] Miller, J.T., Horrocks, H.C. and Dagen, H., "Spectrometric Oil Analysis Program ( SOAP ) Evaluation Criteria ", ARINC Research Corporation Publication No. 1041-01-1-1310, Nov.1974.

[5] Tyler, J.C., and Cuellar, J.P.Jr., " Refinement and Updating of SOAP Wearmetal Evaluation Criteria and Equipment Sampling Intervals ", Southwest Research Institute Report No.RS-655, Aug.1977.

## APPENDIX

In this appendix we shall sketch in the reasoning and computations involved in the prinicipal components approach to determining the sampling interval. It is assumed that $k$ elements are being monitored for the given engine type; $k$ of course would vary with the type of engine. A single record for the given engine consists of the $k$ metallic contaminant readings observed, together with the flight time value at which they were observed. The computations of the principal component to be described employ the previous $n$ records for the given engine, not including the current most recent record for the engine. The current record will be referred to as the $(n+1)^{st}$ record in this scheme.

Let $y_{ij}$ represent the $i^{th}$ reading for element $j$, where $i = 1, 2, ..., n$, and $j = 1, 2, ..., k$. The values of the useage variable (e. g. the flight times) will be denoted by $t_1, t_2, ..., t_n$. The first step required is the computation of (a constant times) the covariance matrix for the observed contaminant readings. This is a $k \times k$ matrix whose $j^{th}$ diagonal element is

$$\sum_i (y_{ij} - \bar{y}_j)^2,$$

and whose $ij^{th}$ off diagonal element is

$$\sum_k (y_{ki} - \bar{y}_i)(y_{kj} - \bar{y}_j),$$

where $\bar{y}_j$ is the average of the $n$ readings for element $j$. Apart from a constant, the diagonal elements are the variances of the readings from the $n$ samples for the $k$ elements and the off–diagonal elements are the covariances between pairs of elements. This matrix is symmetric and in general nonsingular. This implies that it will have $k$ positive characteristic roots, each with a corresponding characteristic vector. The characteristic vector, normalized to have length one, which is associated with the largest characteristic vector is called the first principal component of the matrix. It indicates the direction in the $k$–dimensional space of the $n$ sample vectors which contains the largest amount of the variance of the observed sample values. It is proposed that this first principal component be used to weight the $k$ element values, and that the one–dimensional resulting values be employed to determine the interval.

The first principal component of any matrix, defined above, cannot be expressed in simple closed form and must be determined numerically; many different routines, easily implemented on a micro–computer, are readily available for performing this computation. The result of the computation is simply a vector of $k$ numbers of length one, i. e. whose sum of squares equals 1. We propose that this vector be "re–normalized" so that its sum, not sum of squares, equals 1. This is suggested so that the resulting weighted sums to be described below will maintain the parts per million (ppm) scale of the original readings. The elements of the resulting vector will be denoted by $c_1, c_2, .., c_k$.

Having computed the "renormalized" first principal component, it is now used to weight the values of the k elements:

$$Y_i = \sum_j c_j y_{ij}$$

for $i = 1, 2, ..., n$. This replaces the original $n$ $k$–dimensional vectors by $n$ numbers. These $n$ numbers, weighted averages over the $k$ elements, are then regressed against the values of the useage variable t; compute the slope

$$b = \frac{\sum_i t_i(Y_i - \bar{Y})}{\sum_i (t_i - \bar{t})^2}$$

and the y intercept

$$a = \bar{Y} - b\bar{t}$$

where $\bar{Y}$ is the average of the weighted contaminant values and $\bar{t}$ is the average of the $t_i$ values. Also compute the estimated stantard deviation of the values about the fitted line

$$s = \sqrt{\frac{\sum_i (Y_i - a - bt_i)^2}{n - 1}},$$

which will be used to judge whether the most recent record (the $(n+1)^{st}$, whose contaminant values are denoted by $r_1, r_2, \ldots, r_k$, $t^*$ denotes the value of the current flight hours) is sufficiently large to suggest that the time to the next sample should be shortened. To do this, the weights derived earlier from the first principal component, $c_1, c_2, \ldots, c_n$, are used to weight the current contaminant values:

$$w = \sum_j c_j r_j.$$

Based on the earlier records, we would expect this value w to be essentially

$$a + bt^*;$$

if w is sufficiently large, one might choose to shorten the interval to the next sample. One rule of this sort would be

a. If $w < a + bt^* + s$, continue sampling at the usual rate.
b. If $w \geq a + bt^* + s$, take the next sample at half the usual time.

There are a number of details about this type of procedure which can only be investigated in a meaningful way by actually employing them with real data. The value of $n$, the number of records to employ, would have to be at least as large as $k$, the number of elements monitored, so that the matrix used to determine the first principal component will in fact be nonsingular; perhaps using $n = k + 1$ would be a reasonable choice, but various different values should be tried. Similarly, the above suggestion, that the sampling interval should be cut in half if the actual observed weighted value exceeds the expected plus s, the standard deviation, is arbitrary and should be looked at with real data. Perhaps this sampling interval should be halved if w exceeds the expected plus $q \times s$, where q could be $\frac{1}{2}$, or $\frac{1}{3}$, or 1.24, etc. Only trials with actual data can suggest a "best" value for a factor like q.

To illustrate this methodology, the following data was recorded for F–100 engine 680123. The data has been augmented with a random increment following the single digit output to mimic the actual readings produced by the Baird–Atomic spectrometer. The record labelled Last is assumed to be the current readings; presented above it are the $n = k + 1 = 5 + 1 = 6$ preceding records for the same engine. The time values are the recorded flight hours at the times the samples were taken.

| | Fe | Ag | Cr | Ni | Ti | Time |
|---|---|---|---|---|---|---|
| | 3.5 | 0.1 | 0.2 | 0.2 | 0.8 | 46 |
| | 3.3 | 0.3 | 0.0 | 0.8 | 0.7 | 49 |
| | 2.6 | 0.3 | 0.4 | 0.2 | 1.4 | 53 |
| | 2.7 | 0.4 | 1.2 | 0.7 | 1.4 | 53 |
| | 2.7 | 0.1 | 0.2 | 0.2 | 0.7 | 55 |
| | 2.8 | 0.2 | 0.2 | 0.3 | 0.2 | 55 |
| Last | 3.3 | 1.2 | 0.2 | 0.6 | 1.1 | 57 |

The rounded matrix of (a constant times) the variances and covariances, which leads to the principal component used for the weights is

$$\begin{bmatrix} 0.693 & -0.077 & -0.373 & 0.090 & -0.273 \\ -0.077 & 0.073 & 0.167 & 0.120 & 0.167 \\ -0.373 & 0.167 & 0.913 & 0.180 & 0.673 \\ 0.090 & 0.120 & 0.180 & 0.380 & 0.100 \\ -0.273 & 0.167 & 0.673 & 0.100 & 1.073 \end{bmatrix}.$$

The (rounded) largest characteristic root of this matrix has value 1.901 and the (rounded) corresponding characteristic vector has components $-0.348$ 0.140 0.632 0.109 0.670. From this vector we get the (rounded) "renormalized" weights $c_1 = -0.289$, $c_2 = 0.117$, $c_3 = 0.525$, $c_4 = 0.091$, $c_5 = 0.557$, which sum to 1. These weights then are used with the original 6 records to evaluate the (rounded) values $Y_1 = -0.431$, $Y_2 = -0.456$, $Y_3 = 0.291$, $Y_4 = 0.739$, $Y_5 = -0.255$, $Y_6 = -0.542$, which are regressed against the time values $t_1 = 46$, $t_2 = 49$, $t_3 = 53$, $t_4 = 53$, $t_5 = 55$, $t_6 = 55$, yielding the least squares line with y–intercept $a = -2.093$, and slope $b = 0.038$. The standard deviation about this line is $s = 0.493$. For the current (Last) record, then, the value expected is $a + 57b = 0.089$. The value observed (using the weighted average of the current Last contaminant readings) is $-0.041$, which is in fact below what would be expected. Thus using a rule of the sort mentioned earlier, the decision would be to continue sampling at the same rate.

INITIAL DISTRIBUTION LIST

DIRECTOR (2)
DEFENSE TECH. INFORMATION
CENTER, CAMERON STATION
ALEXANDRIA, VA 22314

DIRECTOR OF RESEARCH ADMIN.
CODE 012
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943

LIBRARY (2)
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943

Center for Naval Analysis
4401 Ford Ave.
Alexandria, VA 22302-0268

DEPARTMENT OF MATHEMATICS
CODE 53
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943

PROF.TOKE JAYACHANDRAN (25)
CODE 53Jy
DEPARTMENT OF MATHEMATICS
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943

Commanding Officer
Kelly Air Force Base
(SA-ALC/MME1)
Texas, 79241